

(19)日本国特許庁 (J P)

(12)公開特許公報 (A)

(11)特許出願公開番号

特開平7-105233

(43)公開日 平成7年 (1995) 4月21日

(51) Int. Cl. ⁶	識別記号	庁内整理番号	F I	技術表示箇所
G 0 6 F 17/30		9194-5L	G 0 6 F 15/403	3 3 0 C
9/44	5 7 0 C	9198-5B	15/40	3 8 0 A

審査請求 未請求 請求項の数 8 O L (全 7 頁)

(21)出願番号 特願平5-250481

(22)出願日 平成5年 (1993) 10月6日

(71)出願人 000005821

松下電器産業株式会社

大阪府門真市大字門真1006番地

(72)発明者 水谷 研治

大阪府門真市大字門真1006番地 松下電器
産業株式会社内

(72)発明者 平井 誠

大阪府門真市大字門真1006番地 松下電器
産業株式会社内

(72)発明者 松島 宏司

大阪府門真市大字門真1006番地 松下電器
産業株式会社内

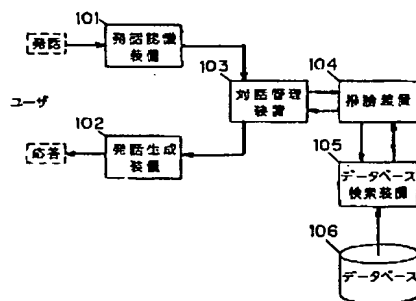
(74)代理人 弁理士 小鍛冶 明 (外 2 名)

(54)【発明の名称】 情報検索装置

(57)【要約】

【目的】 ユーザの自由な発話による検索が可能な情報検索装置を提供すること。

【構成】 ユーザの発話は発話認識装置 101 に入力され、形態素解析、構文解析、意味解析がなされて発話の解釈が対話管理装置 103 に出力される。対話管理装置 103 は対話の文脈を参照して省略表現、照応表現の同定を行い、推論装置 104 に結果を出力する。データベース検索装置 105 は推論装置 104 から出力された対話の文脈に基づきデータベース 106 からデータを検索して推論装置 104 に出力する。推論装置 104 はユーザの意図を満足するデータが得られているかどうかを判断し、対話管理装置 103 にユーザに対する発話の命題表現を出力する。対話管理装置 103 はそれを入力として対話の文脈に応じて修正し、発話生成装置 102 に出力する。発話生成装置 102 は入力された命題表現から自然言語による文を生成し、ユーザへの発話として出力する。



【特許請求の範囲】

【請求項1】 ユーザの発話を入力として依存構造を出力する発話認識装置と、命題を入力として自然言語表現をユーザに出力する発話生成装置と、対話の内容を入力としてデータベースからデータを検索して出力するデータベース検索装置と、前記対話の内容と前記データベース検索装置の出力を解析して次にユーザに発話すべき内容を決定して出力する推論装置と、前記発話認識装置の出力を入力として文脈解析を行って前記推論装置に出力し、かつ前記推論装置の出力を入力として対話の文脈に基づく命題を生成して前記発話生成装置に出力する対話管理装置によって構成されることを特徴とする情報検索装置。

【請求項2】 データベース検索装置は推論装置が出力する対話の内容をデータベースのデータに対する制約の記号表現に変換する制約変換装置と、前記データベースの各データについてその特徴を抽象化した記号表現に変換する索引変換装置と、前記制約変換装置の出力と前記索引変換装置の出力を比較して前記制約の記号表現に対して類似度の高い前記記号表現を持つ前記データを選択する照合装置を有することを特徴とする請求項1記載の情報検索装置。

【請求項3】 制約変換装置は対話の内容を入力としてデータベースのデータに対して制約となる部分を抽出し、前記制約となる部分を依存構造表現で記述することを特徴とする請求項1記載の情報検索装置。

【請求項4】 索引変換装置はデータベースの各データに付随する自然言語の文章を入力として依存構造表現で記述することを特徴とする請求項1記載の情報検索装置。

【請求項5】 照合装置は制約変換装置が出力する依存構造表現と索引変換装置が出力する依存構造表現を比較して構造的かつ意味的に類似度の高いデータを選択するデータ選択装置と、前記データの選択の妥当性を評価する検索結果評価装置を有することを特徴とする請求項1記載の情報検索装置。

【請求項6】 データ選択装置は制約の依存構造表現と索引の依存構造表現の表層的な特徴から高速にデータの絞り込みを行う表層検索装置と、ユーザの知識と概念辞書を用いて意味的な絞り込みをする深層検索装置を有することを特徴とする請求項1記載の情報検索装置。

【請求項7】 検索結果評価装置は制約に対する各データの距離の分布によってデータ選択装置の出力の妥当性を評価することを特徴とする請求項1記載の情報検索装置。

【請求項8】 推論装置は対話管理装置が出力する対話の内容と、データベース検索装置が出力する検索データに対する評価から検索結果に対するユーザの満足度を推定し、ユーザに対する発話の内容を決定するルールを有することを特徴とする請求項1記載の情報検索装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】 本発明はユーザの要求に対して情報を提供する装置に関するものであり、特にデータベース検索装置とコンサルテーション・システムに関するものである。

【0002】

【従来の技術】 近年、CPUの性能の向上や光磁気ディスクなどの2次記憶装置の大容量低価格化により、データベースが提供するデータは質量ともに充実し、また、データの識別子を入力すれば高速に引き出すことが可能になった。しかし、データベースの使い勝手は旧態依然であり、システムの多くは複雑なキーワード論理式やメニュー、SQLをインターフェースとしている。ユーザがシステムに慣れるまでの負担が大きく、また慣れたとしても今度は自分が欲するデータを容易に検索することができない。

【0003】 このような背景からデータベース・システムに対する自然言語によるユーザ・インターフェース技術の開発が行われてきた（例えば、特開平2-201575号公報）。この技術では自然言語で入力されるユーザの要求を解析し、データベース検索に失敗したときはその失敗原因の説明をしたり、自動的に制約の緩和を行って代替データを提示するものである。

【0004】

【発明が解決しようとする課題】 前述の技術は検索対象となるデータベースを関係データベースに限定している。関係データベースにおいては属性の値が数値もしくは定義域が設定された値であるので、ユーザとの対話において特定の属性に注目すればデータ間の差異を計算することが容易である。ユーザの意図と検索結果が一致しなかったときはその差異を用いてユーザに検索失敗の原因を説明をしたり、差異の小さいデータを検索して代替データとして提案することができる。

【0005】 しかし、関係データベースよりも一般的なデータベース、例えばテキスト・データベースなどでは、文書のデータ間の差異が計算できない。各データに付随するキーワード列を比較する方法ではその差異の解釈が困難であり、前述の技術を応用することができない。

【0006】 本発明は一般的なデータベースについて、ユーザがシステムとの自由な自然言語対話を行いながらユーザの意図するデータを提供することを目的とする。

【0007】

【課題を解決するための手段】 データベース内の各データについて、その内容をよく表現する文章の依存構造を解析し、その構造を索引ファイルとして保持する。対話の内容からユーザの検索意図を抽出し、同様に依存構造で表現する。そして依存構造間の距離を定義する。

【0008】

【作用】 インデックス・ファイルとして依存構造表現を利用することにより、キーワードよりも正確にテキスト

データの内容を表示することができる。また、ユーザの検索意図も依存構造で表現することにより、キーワード論理式よりも正確に表現される。データの検索は依存構造表現間に定義した距離の測定によって行われ、適当な数のデータがユーザの検索意図から抽出した制約表現に近いものから順に選択される。また距離の測定が不可能、もしくは多くのデータ間の距離が大きく、測定結果にあまり意味がないと判断される場合は、測定を可能にするために必要な不足情報をユーザに問い合わせることで解決することができる。また、ユーザが検索意図の言語化に困っている場合は、データの距離の分布に最も大きな変化を与えるデータの特徴を同定してユーザに知らせることにより、ユーザへの発想支援を行うことができる。

【0009】

【実施例】本発明の構成を示すブロック図を図1に示す。ユーザの発話は発話認識装置101に入力され、形態素解析、構文解析、意味解析がなされて発話の解釈が対話管理装置103に出力される。対話管理装置103は対話の文脈を参照して省略表現、照応表現の同定を行い、十分な解析結果を得た場合は推論装置104に結果を出力する。データベース検索装置105は推論装置104から出力された対話の文脈に基づきデータベース106からデータを検索して推論装置104に出力する。推論装置104はユーザの意図を満足するデータが得られているかどうかを判断し、対話管理装置103にユーザに対する発話の命題表現を出力する。対話管理装置103はそれを受け取って対話の文脈に応じて修正し、発話生成装置102に出力する。発話生成装置102は入力された命題表現から自然言語による文を生成しユーザへの発話として出力する。

【0010】次に本発明の動作を説明する。実施例の1つとして特許の抄録データベースの検索を行うシステムを取り上げる。対話管理装置103はユーザの発話を解析して推論装置104に出力する。例えば、以下の対話について、

システム： どのような特許をお探ですか。

ユーザ： パイプラインの高速化について書かれたものをお願いします。

対話管理装置103はユーザの発話中の「もの」がシステムの発話の「特許」を指すことを同定し、また、「お願いします」のような曖昧な慣用表現を「要求する」という具体的な単語に変換して、図2のような依存構造を出力する。推論装置104は図2の依存構造がユーザの意図を十分に表現するものであると判断し、データベース検索装置105に出力する。

【0011】図3にデータベース検索装置105の内部構成を示す。制約変換装置301は対話の内容からデータベースのデータに対する制約を抽出し、依存構造で記述する。図2の対話の内容からは「パイプラインの高速

化」という制約を抽出して図4に示す依存構造を生成する。索引生成装置302はデータベース106の各データについて、その内容を端的に表現する文を解析してその依存構造をあらかじめ生成しておく。例えば、ある特許の抄録の目的項の文、「単精度の固定小数点の演算回路にプライオリティエンコーダー回路、マルチプレクサ等の回路を付加することにより、効率よく浮動小数点演算を高速にデータ駆動方式パイプライン上で行えるようにする。」については図5に示す依存構造を生成する。照合装置303は制約変換装置301と索引生成装置302から入力される依存構造を比較し、意味的に類似したデータの選択を行う。

【0012】図6に照合装置203の内部構成を示す。データの依存構造と制約の依存構造はデータ検索装置601でその距離が測られ、距離の小さいものから順位付けされて出力される。検索結果評価装置602はデータの距離の分布から検索結果の妥当性を計算する。

【0013】図7にデータ検索装置601の内部構成を示す。表層検索装置702は依存構造の表層的な特徴に着目して高速な粗いデータ検索を行う。深層検索装置701は表層検索装置が検索したデータの中から、概念辞書やユーザの知識を利用して意味的に距離が近いデータの検索を行う。例えば、表層検索装置702は図4の制約の依存構造から「パイプライン」という単語が表層的な制約として有効であると判断し、依存構造の中にその単語を含むデータを選択する。抄録文として、「結合演算処理に係わるサイト及び分散データベースを利用する為の計算機とをを通じてパイプライン処理的に結合演算を実行することにより、結合演算を高速に行う。」および、「パイプラインとは別に部分書込みの為のデータのみを保持するレジスタ群を設けることにより、パイプラインの規模を小さくする。」を持つ特許は、その依存構造がそれぞれ図8と図9のようになるので、表層検索装置702によって選択される。深層検索装置701は概念辞書を用いて図4の構造に対して図5、図8、図9の距離を測定する。距離の近いものは図5、図8、図9の順になるが、どの距離の値も大きい。検索結果評価装置602はデータに対して図4の制約に対する距離の分布を出力し、最も距離が近いのは図5であることを出力する。

【0014】推論装置104は対話管理装置103が出力する対話の内容と、データベース検索装置105が出力する検索データと検索結果評価に基づき、ユーザに対する発話の内容を図10に示す状態遷移図を用いて決定する。例では各データとの距離が大きいので、最も距離が近いデータの図5において図4の制約と類似する構造を抽出し、命題、(ask user (演算する (様式 高速に) (で パイプライン) (を) (方法)))

を対話管理装置103に出力し、次の発話の内容をユー

ザに対する質問とする。対話管理装置103は対話の文脈を参照して、入力された命題を、
(ask user (演算する (様式 高速に) (で パイプライン)))

という命題に変換して発話管理装置102に出力する。発話管理装置102はユーザに対して、
システム: パイプラインで高速に演算するということですか。

という発話を行う。

【0015】 以上のようにユーザが満足したと推論装置104で判断されるまで自然言語対話によるデータの検索が続けられる。

【0016】 なお、本発明の応用範囲は特許抄録のデータベースに限定されず、絵画や音楽のデータベースでも自然言語による文章が各データに付随しているものであれば応用することができる。

【0017】

【発明の効果】 ユーザは検索論理式やキーワードによる検索戦略を考える必要がなく、自然言語による自由な対話を通じて自分が意図するデータを検索することができる。

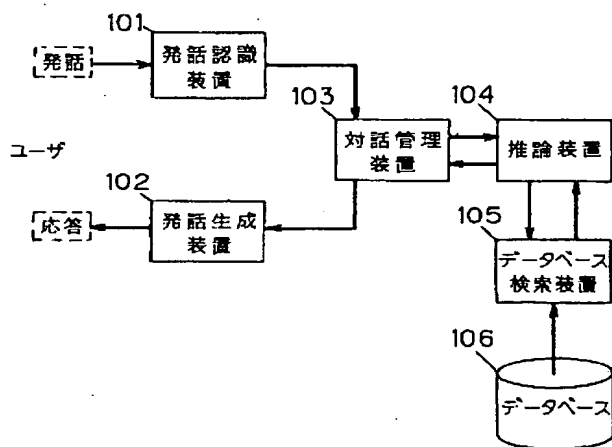
【図面の簡単な説明】

【図1】 本発明の全体の構成を表すブロック図

【図2】 ユーザの要求の依存構造表現の一例を示す図

【図3】 データベース検索装置105の内部構成を表す

【図1】



ブロック図

【図4】 データに対する制約の依存構造表現の一例を示す図

【図5】 特許の抄録の依存構造表現の一例を示す図

【図6】 照合装置303の内部構成を表すブロック図

【図7】 データ検索装置601の内部構成を表すブロック図

【図8】 特許の抄録の依存構造表現の一例を示す図

【図9】 特許の抄録の依存構造表現の一例を示す図

【図10】 推論装置104の動作を表す状態遷移図

【符号の説明】

101 発話認識装置

102 発話生成装置

103 対話管理装置

104 推論装置

105 データベース検索装置

106 データベース

301 制約生成装置

302 索引生成装置

303 照合装置

601 データ検索装置

602 検索結果評価装置

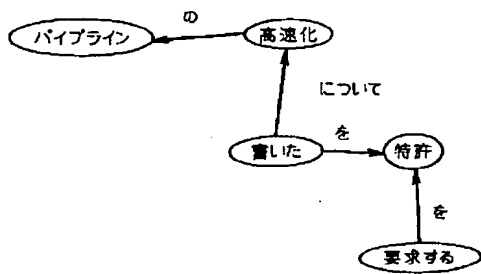
701 深層検索装置

702 表層検索装置

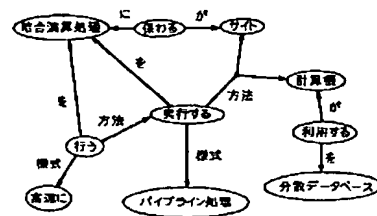
【図4】



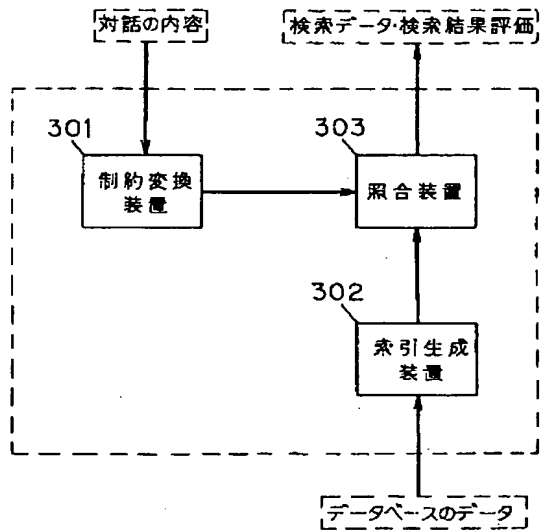
【図2】



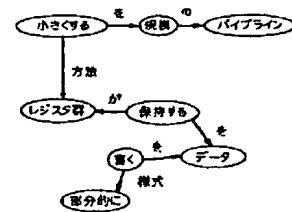
【図8】



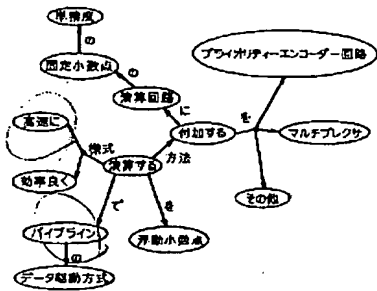
【図3】



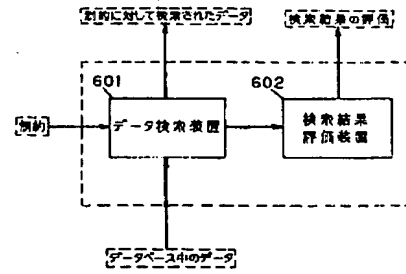
【図9】



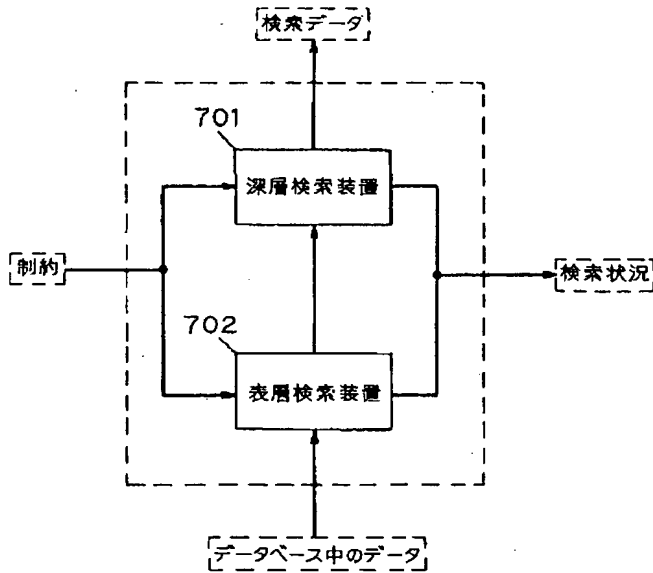
【図5】



【図6】



【図7】



【図10】

